# A Monte Carlo-quantum mechanics study of the solvatochromic shifts of the lowest transition of benzene

Kaline Coutinho[a] and Sylvio Canuto
*Instituto de Física, Universidade de São Paulo, CP 66318, 05315-970 São Paulo, SP, Brazil*

M. C. Zerner
*Quantum Theory Project, University of Florida, Gainesville, Florida 32611*

(Received 31 August 1998; accepted 4 February 2000)

We examine the spectroscopic red shifts that occur when benzene is dissolved in (liquid) benzene, in cyclohexane, in carbon tetrachloride, and in water. For this we develop a mixed classical/quantum model in which uncorrelated structures are obtained from Monte Carlo simulation, and these structures are then used for quantum chemical calculations including the chromophore and all solvent molecules within the first radial distribution maxima. We discuss the effects of different sampling techniques and the inclusion of more, or less, solvent molecules in the quantum chemical supermolecule calculation. We obtain shifts of $-306\,\mathrm{cm}^{-1}$, $-268\,\mathrm{cm}^{-1}$, $-456\,\mathrm{cm}^{-1}$, and $-122\,\mathrm{cm}^{-1}$, in excellent agreement with the experimentally observed shifts of $-332\,\mathrm{cm}^{-1}$, $-308\,\mathrm{cm}^{-1}$, $-458\,\mathrm{cm}^{-1}$, and $-143\,\mathrm{cm}^{-1}$, respectively. We note that the larger shift observed in carbon tetrachloride that is not expected on the basis of larger dielectric constant results from small contributions of the charge transfer type from solvent to solute. © *2000 American Institute of Physics.* [S0021-9606(00)51716-4]

## I. INTRODUCTION

Solvent effects have been the subject of considerable interest.[1–5] In particular, the study of solvent effects in molecular absorption spectroscopy in the visible-UV region has its own motivation and complexities.[6,7] Bayliss has earlier used classical analysis to relate spectral shifts to the index of refraction under the assumption that the solvent influences the shift by a dielectric constant.[8] More recently, the idea of a solute enclosed by a cavity and interacting with the solvent as a polarizable dielectric continuum has been extended in many ways.[9,10] As this reaction field is a response of the medium to a permanent dipole moment of the solute it is not expected to work properly for the cases of a nonpolar solute such as benzene. Another approach that is gaining wide acceptance is the use of a combined molecular mechanics/quantum mechanics (MM/QM) approach[11–15] in which the most important part of the problem is treated by quantum mechanics and the rest by classical force fields. More recently computer simulation methods are becoming a general tool for treating liquid systems. Thus the idea of generating liquid structures of one solute surrounded by solvent molecules is a natural one. However for these large, real supermolecular systems, consisting of hundreds of molecules, a complete quantum mechanical calculation is largely impractical. This is complicated even further because a liquid has many structures (geometrical arrangements of atoms and molecules) equally possible and some sort of ensemble average has to be performed thus implying a very large number of quantum mechanical calculations.

In this paper we address to the solvatochromic shifts of benzene in different solvents. As benzene is neutral and has no permanent dipole moment both in the ground and in the $\pi-\pi^*$ excited states the dipolar self-consistent reaction field is essentially zero and no spectral shift of these bands can be expected using this model. Of course, higher moments can be used but these give smaller contributions. However, the experimental results for the first absorption band of benzene in water, cyclohexane, and carbon tetrachloride, for instance, show clear red shifts as compared to the gas phase results. As we shall see excellent results will be presented in this paper using a combination of Monte Carlo (MC) and quantum mechanical calculations. We essentially use a sequential Monte Carlo simulation-quantum mechanical approach[16] to obtain solute–solvent interactions and solvatochromic shifts. First the Monte Carlo simulation is performed in the solute–solvent liquid. During the simulation molecular structures are obtained and separated. These structures are used in a subsequent quantum mechanical calculation. As a MC simulation involves several thousands of steps for averaging of the thermodynamic properties, and now also of the transition energies, a specific procedure for reducing the total number of structures needed for the quantum calculation is of great importance. We shall show that this is, indeed, possible using the concept of time correlation or statistical efficiency. Therefore, instead of performing a quantum mechanical calculation on every structure generated by the Monte Carlo simulation these calculations are performed only on some few uncorrelated structures without compromising the final average result.[17] Also, as each structure obtained by the simulation contains hundreds of solvent molecules, a quantum mechanical treatment of the solvent requires a judicious reduction of the total number of solvent molecules included.

[a]Permanent address: Universidade de Mogi das Cruzes, CP 411, 08701-970 Mogi das Cruzes, SP, Brazil.

 

TABLE I. Information on the simulated systems of this study: the number of molecules $N$, the density $\rho$ in g/cm$^3$, the cutoff radius $r_c$ in Å and the average number of molecules inside the cutoff radius $N_{\text{eff}}$.

| System | $N$ | $\rho$ | $r_c$ | $N_{\text{eff}}$ |
|---|---|---|---|---|
| Liquid $C_6H_6$ | 1+124 | 0.8990 | 13.11 | 65 |
| $C_6H_6$ in $C_6H_{12}$ | 1+124 | 0.7785 | 14.10 | 65 |
| $C_6H_6$ in $CCl_4$ | 1+343 | 1.5867 | 19.04 | 175 |
| $C_6H_6$ in $H_2O$ | 1+343 | 0.9966 | 10.92 | 180 |

We therefore analyze the results for the radial distribution function that defines the solvation shells. All molecules within a certain shell are explicitly included.

## II. MONTE CARLO SIMULATION

Monte Carlo simulations are carried out for four systems: liquid benzene, benzene in cyclohexane, in carbon tetrachloride and in water. Standard procedures are used for the Metropolis sampling in the canonical ($NVT$) ensemble.[18,19] All the calculations are performed at $T=25$ °C. The volume of the cubic box is determined by the experimental density of the liquid and the number of molecules used in the simulation. As usual, periodic boundary conditions, the image method and a cutoff radius are used. For all molecules included in the simulation, each one interacts effectively with $N_{\text{eff}}$ molecules. In Table I, the number of the molecules $N$, the density $\rho$, the cutoff radius $r_c$, and the $N_{\text{eff}}$ are shown for all simulated systems. All the simulations are performed using the DICE program.[20]

The intermolecular interactions are described by the Lennard-Jones plus Coulomb potential. The interaction energy between molecules $a$ and $b$, $U_{ab}$, is the sum of the individual interactions between all pairs of sites on the two molecules,

$$U_{ab}=\sum_{i}^{\text{on}a}\sum_{j}^{\text{on}b}\left\{4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12}-\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right]+\frac{q_iq_je^2}{(4\pi\varepsilon_0)r}\right\}, \quad (1)$$

where $r_{ij}$ is the distance between site $i$ (on molecule $a$) and site $j$ (on molecule $b$), $\varepsilon_{ij}$ and $\sigma_{ij}$ are combinations of the Lennard-Jones parameters of sites $i$ and $j$, $q_i$ is the charge of the site $i$ and $e^2/(4\pi\varepsilon_0)=331.9684$ Å kcal/mol. We use the potential parameters $q_i$, $\varepsilon_i$, and $\sigma_i$ (shown in Table II) as suggested in the literature for benzene and cyclohexane,[21] for carbon tetrachloride[22] and for water.[23] In all simulated systems, the molecules have rigid geometries. Distortions of the benzene structure would lead to considerably enhanced intensity of the forbidden $^1B_{2u}$ transition.[24] However, preliminary investigation indicates that the position of the band maximum will not be much effected. The benzene molecules are in the $D_{6h}$ structure with $r_{CC}=1.400$ Å, $r_{CH}=1.088$ Å and all angles $\Theta=120$ °C. The cyclohexane molecules are in the $C_{2v}$ chair structure with $r_{CC}=1.53$ Å, $r_{CH}=1.088$ Å, $\Theta(\text{CCC})=112°$ and $\Theta(\text{HCH})=110.29°$ (as is well known the chair structure composes 99% of the liquid phase[25]). The carbon tetrachloride molecules are in the $T_d$ structure with $r_{\text{CCl}}=1.766$ Å and $\Theta(\text{ClCCl})=109.47°$ and the water molecules are in the $C_{2v}$ structure with $r_{\text{OH}}=1.000$ Å and $\Theta(\text{HOH})=109.47°$.

TABLE II. The potential parameters [see Eq. (1)] of the Monte Carlo simulation. $q_i$ in elementary charge unit, $\varepsilon_i$ in kcal/mol, and $\sigma_i$ in Å.

| Site | $q_i$ | $\varepsilon_i$ | $\sigma_i$ |
|---|---|---|---|
| C of $C_6H_6$[a] | 0.000 | 0.110 | 3.750 |
| H | 0.000 | 0.000 | 0.000 |
| C of $C_6H_{12}$[a] | 0.000 | 0.118 | 3.905 |
| H | 0.000 | 0.000 | 0.000 |
| C of $CCl_4$[b] | 0.000 | 0.102 | 4.600 |
| Cl | 0.000 | 0.203 | 3.500 |
| O of $H_2O$[c] | −0.820 | 0.150 | 3.175 |
| H | 0.410 | 0.000 | 0.000 |

[a]Reference 21.
[b]Reference 22.
[c]Reference 23.

The initial configuration is generated randomly, considering the position and the orientation of each molecule. A new configuration, or one MC step is generated by selecting all molecules sequentially and trying to translate it in all the Cartesian directions and also rotate it around a randomly chosen axis.[19] The maximum allowed displacement of the molecules is autoadjusted after 50 MC steps to give an acceptance rate of new configurations around 50%. The maximum rotation angle is fixed during the simulation in $\delta\Theta=\pm 15°$. Here one MC step is generated after *all* molecules sequentially attempt to move. This MC step corresponds in the more usual definition[23] to $N$ MC steps, where $N$ is the total number of molecules in the simulation. The full simulation involves a thermalization stage of ~5000 MC steps followed by an averaging stage of ~50 000 MC steps. (For example, in the cyclohexane and carbon tetrachloride simulations these 50 000 steps translates into 343 molecules $\times 50\,000=1.7\times10^7$ individual moves, adopted in such a fashion that about 50% are accepted.) The averaging step has also been checked to insure that the average shifts we calculate and report do not drift as a function of sampling time. The structures that we generate from this procedure and these force fields seem reasonable, and have been checked in some detail in the case of liquid benzene against the observed radial distribution function.[16] In this reference we noted that the experimental structure of the solid yielded calculated shifts of benzene of $-260\,\text{cm}^{-1}$ compared to the experimentally observed value of $-260\,\text{cm}^{-1}$. This suggests that the quantum mechanics we are using is sufficient. Assuming the reported liquid structure yielded shifts between $-60\,\text{cm}^{-1}$ and $-100\,\text{cm}^{-1}$, small when compared to the experimental value of $-320\,\text{cm}^{-1}$. The full liquid simulation, as described below and in that reference, leads to a less structured liquid, but still very reasonable agreement with the observed radial distribution function, and an excellent predicted spectroscopic shift.[16]

The average value of the energy is given as the average over a chain of size $T$ of instantaneous values,

$$\langle U\rangle=\frac{1}{T}\sum_{i}^{T}U_i \quad (2)$$

and the error caused by the finite size of the simulation can be obtained as

$$\sigma(\langle U\rangle)=\sqrt{\frac{1}{T}\langle \delta U^2\rangle}, \tag{3}$$

where

$$\langle \delta U^2\rangle=\frac{1}{T}\sum_{i}^{T}(U_i-\langle U\rangle)^2=\langle U^2\rangle-\langle U\rangle^2 \tag{4}$$

is the variance of $U$. Clearly, if $T$ approaches infinity the error approaches zero.

The number $T$ of MC configurations necessary for the ensemble average can be reduced dramatically over the entire number of MC steps without modifying the calculated average energy. This is a consequence of the high statistical correlation between successive MC configurations generated from simulations of dense systems such as liquids. The changes between successive configurations are small and cause high statistical correlation between them. Using statistically correlated configurations is of course a waste that does not contribute effectively to the average. An analysis of the efficiency of the simulation shows how many MC steps are necessary to statistically uncorrelate the configuration. This can be done by calculating the time (or step) correlation function.[19]

After the simulation, the time correlation function, $C(t)$, can be calculated according to the following equation:

$$C(t)=\frac{\langle \delta U_i \delta U_{i+t}\rangle}{\langle \delta U^2\rangle}, \tag{5}$$

where $t$ here means MC step, $U_i$ is the energy of the $i$th configuration, $U_{i+t}$ is the energy of the configuration generated $t$ steps later and $\langle \delta U_i \delta U_{i+t}\rangle=\langle U_i U_{i+t}\rangle-\langle U_i\rangle\langle U_{i+t}\rangle$. For Markovian processes the correlation function satisfies an exponential decay[26,27]

$$C(t)=\sum_{i}^{n} c_i e^{-t/\tau_i}, \tag{6}$$

where the correlation time $\tau$ is

$$\tau=\int_0^\infty C(t)dt. \tag{7}$$

In general, configurations separated by $t>2\tau$ have a statistical correlation that is less than 15%. Therefore the number of MC steps necessary to obtain configurations, that are statistically ''uncorrelated,'' is $2\tau$. A concept closely related to the correlation time is the statistical inefficiency.[28] It is based on the mean-square deviations of averages taken over blocks of a chain of successive configurations. Dividing the chain of configuration in $n_b$ blocks with $t_b$ successive configurations, where $T=n_b t_b$, then the average taken over the $b$th block is

$$\langle U\rangle_b=\frac{1}{t_b}\sum_{i=b_s}^{b_e} U_i, \tag{8}$$

where $b_s=(b-1)n_b+1$, $b_e=bt_b$ and $b$ can assume values from 1 to $n_b$. The mean-square deviation of averages taken over the blocks is

$$\langle \delta U^2\rangle_b=\frac{1}{n_b}\sum_{b=1}^{n_b}(\langle U\rangle_b-\langle U\rangle)^2 \tag{9}$$

and the statistical inefficiency $s$ is

$$s=\lim_{t_b\to\infty}\frac{t_b\langle \delta U^2\rangle_b}{\langle \delta U^2\rangle}=\lim_{t_b\to\infty} S(t_b), \tag{10}$$

where $\langle U\rangle$ is the average and $\langle \delta U^2\rangle$ is the variance over the chain of the whole simulation as shown in Eqs. (2) and (4), while the $\langle U\rangle_b$ is the average and $\langle \delta U^2\rangle_b$ is the variance over the blocks. For chains of correlated configurations, $S(t_b)$ will rise with increasing $t_b$ until it approaches a limiting value $s$. The limiting value of $S(t_b)$ signifies that the block size $t_b$ has become so large that there is no correlation from block to block. However, for finite chains ($T=n_b t_b$), as the block size $t_b$ increases the number of blocks $n_b$ decreases and it generates an increased error in the calculation of $S(t_b)$ and consequently an imprecision in the estimate of the best value of $s$. Therefore, an analysis of $s$ with respect to the size of the chain $T$ is very important.

For chains of uncorrelated configurations, Eq. (3) is replaced by

$$\sigma(\langle U\rangle)=\sqrt{\frac{1}{l}\langle \delta U^2\rangle}, \tag{11}$$

where $l=T/s$. In this way, one can see that for a finite simulation the error of a property $U$ is the same if one takes $T$ successive configurations or equivalently $T/s$ configurations separated by $s$ MC steps. Thus for a given simulation obtaining ensemble averages over $T/s$ uncorrelated configurations is the same as over the entire number of configurations generated. For Markovian process $s\approx 2\tau$. These two procedures will be used independently in a complementary way to obtain the best statistics for uncorrelated structures of a finite simulation.

The quantum chemistry calculations are of the SCF type followed by configuration interaction over singly excited configuration state functions (CIS). This is the level of theory at which the INDO/S Hamiltonian was parametrized.[29] The low energy spectrum of benzene is dominated by excitations from the homo and homo$-1$ $e_{1g}(\pi)$ molecular orbitals to the lumo and lumo$+1$ $e_{2u}(\pi^*)$ molecular orbitals, where homo refers to the highest occupied molecular orbital and lumo to the lowest unoccupied molecular orbital. This gives rise to states of $B_{2u}$, $B_{1u}$, and $E_{1u}$ symmetry. The lowest lying of these states is of interest here, the $^1B_{2u}$, and this state is remarkably insensitive to the size of the CIS used. For example, a CIS consisting only of the four frontier orbitals (2 down, 2 up, $2\times2$) yields, for a gas phase calculation, 38 580 $cm^{-1}$ for this transition, the same value obtained for a $9\times12$ CIS. The experimental value is 38 390 $cm^{-1}$. The $9\times12$ CIS is chosen so as not to break the symmetry of the resulting states. This same insensitivity is not apparent for the higher lying states. For example, the two different CI's yield for the second state, the $^1B_{1u}$, 49 515 $cm^{-1}$ and 47 815 $cm^{-1}$, respectively. The experimental value for this transition is 48 960 $cm^{-1}$.

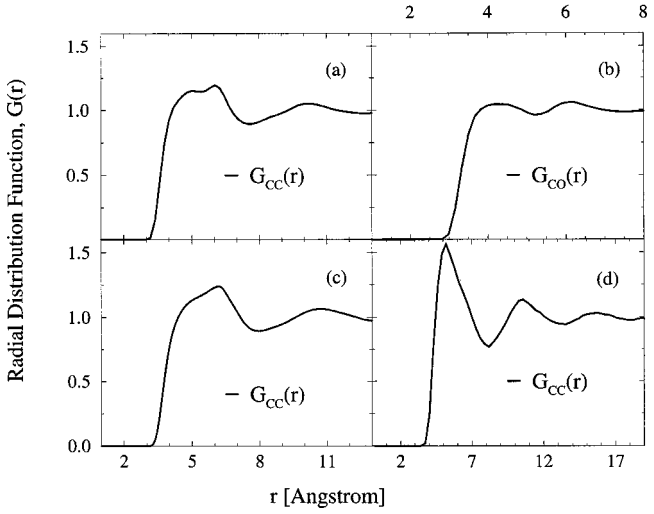


FIG. 1. The calculated radial distribution function (RDF) between benzene and benzene, water, cyclohexane, and carbon tetrachloride. In this figure are shown the RDF between the carbon atoms of benzene and the (a) carbon atoms of benzene; (b) oxygen atom of water; (c) carbon atoms of cyclohexane; and (d) carbon atom of $CCl_4$. Relevant numerical information is given in Table III.

This lack of sensitivity is also apparent in the supermolecule calculations described below. Benzene in water was examined using 2×2, 10×10, 20×20, and 28×28 CIS calculations. There are no differences in the shifts calculated within the accuracy of the statistics reported below. A similar investigation was made for cyclohexane, and again there is no significant sensitivity to the size of the CI. For $CCl_4$ a 16×16 CI is performed to insure that the benzene frontier orbitals are included in the CI, as the molecular orbitals of the solvent are interspersed with those of benzene. For benzene in benzene a 28×28 CI is used to include the $e_{1g}(\pi)$ and $e_{2u}(\pi^*)$ orbitals of all 14 benzenes of the supermolecule.

## III. RESULTS AND DISCUSSIONS

A full quantum mechanical calculation of the energy of a supermolecular system, composed of hundred molecules, over thousands of MC steps is, of course, not possible. The alternative we suggest is to perform quantum mechanical calculations after the simulation but using only a few selected number of solvent molecules and a selected number of MC structures. The number of solvent molecules included in the calculation is obtained from the analysis of the radial distribution function, using all molecules surrounding the solute up to a certain solvation shell. The value of the energy is thus performed by ensemble average over the statistically *uncorrelated* structures generated in the simulations. This procedure reduces dramatically the number of configurations used in the quantum mechanical calculations, when compared with the entire number of MC steps, without modifying the calculated average energy.[16,17]

In Fig. 1 we show the pairwise radial distribution function (RDF) of benzene in the four solvents considered here. Figure 1(a) shows the carbon–carbon RDF of homogeneous liquid benzene. This RDF is in good agreement with the experimental result of Narten.[30,31] It shows a first maximum at 6.0 Å. The first minimum at 7.6 Å defines the limits of the first solvation shell. Integration of the RDF up to 7.6 Å shows that each benzene has a total of 13 benzene neighbors. Thus within the limits of the first solvation shell the supermolecule is composed of 1+13 benzenes. Figure 1(b) shows the RDF of benzene in water. The first maximum is located at 4.4 Å and the first minimum at 5.3 Å. Integration of the first solvation shell shows that the central benzene is surrounded by 18 water molecules. Similarly, in Figs. 1(c) and 1(d), the integration of the first solvation shell, until 7.9 and 8.2 Å, respectively, shows that the central benzene is surrounded by 11 cyclohexane and 14 $CCl_4$ molecules. Table III summarizes the information obtained from the RDF shown in Fig. 1. We now analyze how the calculated energy shifts of the first $\pi-\pi^*$ transition of benzene vary with the number of solvent molecules included in the quantum calculation. The calculated energy shift stabilizes when the number of solvent molecules reaches the first solvation shell, obtained here by integration of the first peak of the RDF. We illustrate this analysis in Fig. 2 that shows the calculated energy shift over just one MC configuration for each system. A similar behavior is obtained for all MC structures but only one result is shown to avoid excessive congestion in the figure. In the following QM calculations we will use different supermolecular sizes for the different solvents. Each supermolecule will be composed of the central benzene solute

TABLE III. Structural information obtained from the radial distribution function. Distances in Å. $N_s$ is the coordination number (see text and Fig. 1).

| System | RDF | 1st max. | 1st shell | $N_s$ |
|---|---|---|---|---|
| Liquid $C_6H_6$ | $G_{CC}$ | 6.0 | 7.6 | 13 |
| $C_6H_6$ in $C_6H_{12}$ | $G_{CC}$ | 6.2 | 7.9 | 11 |
| $C_6H_6$ in $CCl_4$ | $G_{CC}$ | 5.3 | 8.2 | 14 |
| $C_6H_6$ in $H_2O$ | $G_{CO}$ | 4.4 | 5.3 | 18 |

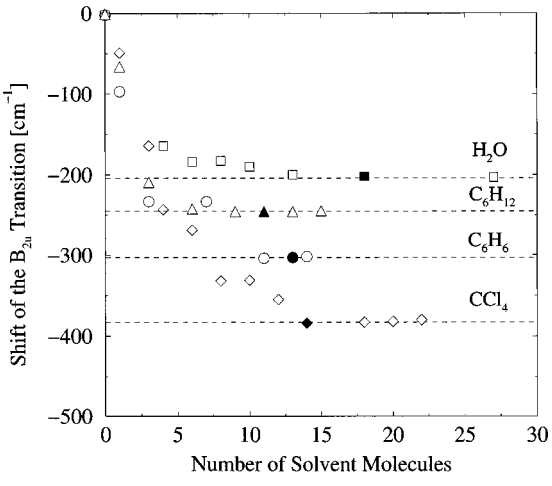


FIG. 2. The calculated energy (red) shift of the $B_{2u}(\pi-\pi^*)$ transition as a function of the number of solvent molecules included in the quantum calculation. The symbols for the solvents are: square for water, triangle for cyclohexane, circle for benzene, and diamond for $CCl_4$. Note saturation after the filled symbols that correspond to the first solvation shell.

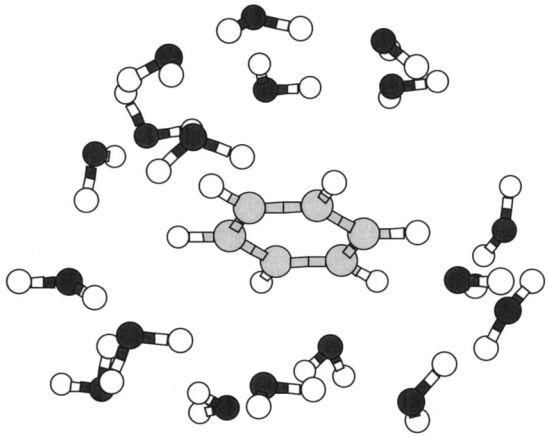

FIG. 3. Typical supermolecule used in the QM calculations. This illustration shows the benzene surrounded by 18 first-neighbor water molecules corresponding to the first solvation shell.

surrounded by 18 molecules in the water solvent, 11 molecules in cyclohexane solvent, 13 molecules in the liquid benzene, and 14 molecules in $CCl_4$ solvent. Figure 3 shows one picture of the supermolecule used in the QM calculation of benzene in water.

Figure 4 shows the calculated results for the time correlation function and the statistical inefficiency for the case of benzene in $CCl_4$. In the left side, we systematically use Eq. (10) to calculate the $S(t_b)$ for chains with different sizes. We obtain the statistical inefficiency, $s=470\pm70$ MC steps. In the right side, we use Eq. (5) to calculate $C(t)$, that was fitted by two exponential functions as in Ref. 32 (see in Fig. 4). The integration of exponential functions gives the correlation time $\tau=230$ MC steps. Now, based on calculated $\tau=230$ and $s=470$, we select from the 25 000 saved configurations of benzene in $CCl_4$, a set of 50 configurations separated by 500 MC steps, e.g., a set of 50 uncorrelated configurations. Therefore, among those MC structures obtained from the simulation only those separated by $\sim$500 MC steps

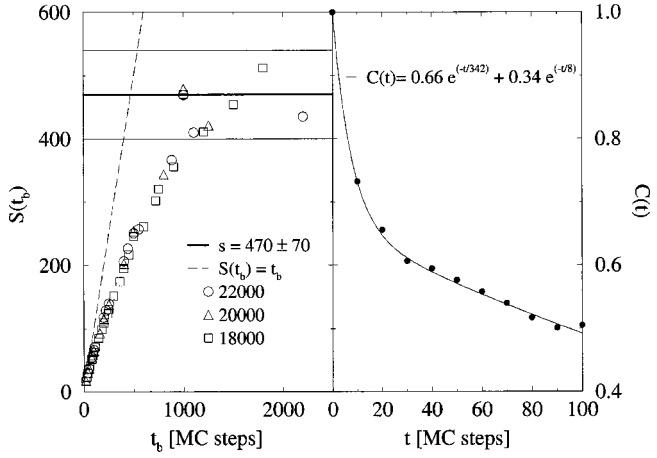


FIG. 4. Statistical inefficiency (left) and time correlation function (right) calculated for the case of benzene in $CCl_4$. On the left side, the behavior of $S(t_b)$ vs $t_b$ for chains with different sizes. The boldface lines represent the asymptotic behavior of correlated blocks (dashed) and uncorrelated blocks (solid). On the right-hand side, the calculated correlation function (dots) are fitted to the exponential form shown. $\tau$ is obtained using Eq. (7) (see text).

TABLE IV. Calculated solvatochromic shift of the $^1A_{1g}-^1B_{2u}(\pi-\pi^*)$transition of benzene in different solvents. $N$ (one solute $+$ $N_s$ solvent) is the number of molecules used in the QM calculations, $l$ is the number of selected configurations to perform QM calculations, $s$ is the number of MC steps between the selected configurations, and $\langle\Delta E\rangle_l$ is the calculated average solvatochromic shift (in $cm^{-1}$). Uncertainty in $\langle\Delta E\rangle_l$ is obtained from Eq. (11). Experimental results are from Ref. 35.

| System | $N$ | $l$ | $s$ | $\langle\Delta E\rangle_l$ | $\Delta E_{exp}$ |
|---|---|---|---|---|---|
| Liquid $C_6H_6$ | 1+13 | 60 | 800 | $-306\pm15$ | $-332\pm30$ |
| $C_6H_6$ in $C_6H_{12}$ | 1+11 | 100 | 500 | $-268\pm20$ | $-308\pm20$ |
| $C_6H_6$ in $CCl_4$ | 1+14 | 50 | 500 | $-456\pm38$ | $-458\pm20$ |
| $C_6H_6$ in $H_2O$ | 1+18 | 100 | 800 | $-122\pm7$ | $-143\pm20$ |

will give statistically new information. Hence, in the QM calculation of the supermolecular structures we will not use all $T$ structures generated but only $l=T/s$ structures separated by $s$ MC steps. It is our experience[17] that the results do not depend on the particular set of MC structures used as long as they are separated by $s$ steps in a relatively long simulation. For the other systems, using the same analysis we selected 60 structures of liquid benzene separated by 800 MC steps, 100 structures of benzene in water separated by 800 MC steps and 100 structures of benzene in cyclohexane separated by 500 MC steps. These were enough to give stable value for the average shift of the transition energy $\langle\Delta E\rangle_l$.

All QM calculations are made at the semiempirical level using the ZINDO program.[33] The energy shift $\Delta E$ is obtained for all supermolecules using singly excited CI with spectroscopic parametrization as described above (INDO/S).[29,34] The ground state wave function is obtained from a self-consistent field (SCF) calculation taking into account antisymmetry over the entire solute–solvent system. Each supermolecule is composed of the solute benzene and the $N_s$ nearest neighbors obtained after integration of the first peak of the RDF. The total number of supermolecules used for averaging the solvatochromic shift for each solvent is obtained by dividing the total number of MC steps used in the simulation by the statistical inefficiency $s\approx2\tau$. This information is collected in Table IV that also reports our final results for the solvent shifts of the first absorption band of benzene. As can be seen, the final calculated solvatochromic shifts $\langle\Delta E\rangle_l$ are in excellent agreement with the experimental results in all cases considered. For the case of benzene in $CCl_4$, for instance, 50 QM calculations were performed in supermolecules separated by 500 MC steps, each one composed of one benzene surrounded by 14 $CCl_4$ molecules. The final average result for the solvatochromic shift is $-456\,cm^{-1}$, in excellent agreement with the experimental result of $-458\,cm^{-1}$.[35] It should be noted that some supermolecules give $\Delta E$ values that deviates greatly from the experimental result and that would lead to spectral line broadening, but the final average $\langle\Delta E\rangle_l$ is in excellent agreement with experiment. The numbers in Table IV give a clear indication that single cluster model cannot represent the true situation of a disordered liquid. Also, the size of the supermolecular system seems to include both the short and the long range solute–solvent interaction, although this is un-

likely to be true for charged or highly polar systems. For the case of benzene in water, for instance, all water molecules within the range of 6.0 Å from the solute are explicitly included. It is expected that the relevant induction-polarization effects are included. In the specific case of the $CCl_4$ we note small contributions of the charge-transfer type from $CCl_4$ to benzene. At the orbital level the charge transfer in $CCl_4$ originates with molecular orbitals of the solvent interspersed with those of benzene, thus appearing also $\pi - p^*$ contributions. The $e_{1g}(\pi)$ and $e_{2u}(\pi^*)$ molecular orbitals that give rise to the $^1B_{2u}$ state, delocalize little over the solvent. For water they are 90% localized in the benzene whereas for cyclohexane and $CCl_4$ it is around 80% in the benzene with 20% of delocalization over the solvent.

## IV. SUMMARY AND CONCLUSION

A sequential procedure using Monte Carlo simulations and quantum mechanics is used to calculate solvatochromic shifts of benzene in different solvents, a case in which dispersion contributions should dominate and lead to red shifts. Quantum-mechanical supermolecular calculations are performed on structures generated by the simulation. The total shift is obtained as an average over many uncorrelated structures. The size of the supermolecule upon which the quantum chemistry is performed is obtained after analysis of the radial distribution function. The number of structures used to calculate the average energy is obtained after consideration of the time correlation function and the statistical inefficiency. Only structures that are statistically uncorrelated are used. The calculated values for the solvatochromism of benzene in benzene, water, cyclohexane, and carbon tetrachloride are found to be in excellent agreement with the experimental result.

The solvatochromic shifts of benzene, a nonpolar solute, in nonpolar solvents is believed to be mostly dispersive in nature. Indeed, we estimate that the quadrupolar contribution to the shift of benzene in cyclohexane using self-consistent reaction field theory[9] to be $12\,cm^{-1}$ to the blue. Using the model for dispersion described in Ref. 36 we can estimate the shift due to dispersion of benzene in cyclohexane to be 540 $cm^{-1}$, a net red shift of 530 $cm^{-1}$, nearly 200 $cm^{-1}$ larger than that calculated here, and that observed. This model uses second order perturbation theory from separate calculations on the solute and on the solvent, and requires the assumption of a cavity size. Although the shift is calculated too large, the ground state stabilization (2136 $cm^{-1}$) and the excited state stabilization (2678 $cm^{-1}$) due to dispersion is *smaller* than that advocated in most other models.[37,38] Although the present calculations are considerably more complex than are those using perturbation theory, the present calculations make no assumption about cavity size or shape, and are clearly quantitatively accurate, and hopefully can be used to examine in more detail more qualitative models.

It should be noted that it is quite remarkable that the present supermolecule calculations reproduce the spectroscopic shifts of nonpolar benzene in nonpolar solvents using a singles only CI method. The dispersion of the ground state requires double excitations, and dispersion of excited states dominated by single excitations, as in the $^1B_{2u}$ of benzene, requires triply excited configurations, a select group of those that are doubly excited relative to the singles that dominate the description of this state. This select group includes singles localized on the chromophore relative to the dominant singles describing the state [in this case double excitations on the benzene in which one excitation involves the promotion $e_{1g}(\pi)$ to $e_{2u}(\pi^*)$] and singles on the solvent. The observation that a CIS model is successful has led us to prove a rather remarkable theorem that this should be expected: dispersion is included in the energy *differences* calculated in a CIS model, even though it is not included in the energies of the individual states.[39]

This study was initiated to examine dispersion effects on spectroscopy, and led us to studying benzene in various solvents where there is ample experimental information with which to compare, and in which the leading causes of the shifts are dispersion. Self-consistent reaction field approaches are simpler, and seem to predict spectroscopic shifts for polar molecules quite accurately, especially when comparing shifts in various different solvents, where the dispersive shifts are roughly the same regardless of the nature of the solvent.[4,10,38,40] Even in this study, the dispersion shifts between two very different solvents, water and $CCl_4$, is no greater than 300 $cm^{-1}$, while shifts due to the changes in permanent moments between the ground and excited states of polar molecules can be an order of magnitude larger. Shifts caused by changes in permanent moments upon excitation are, in general, easy to understand in the sense that stabilization is proportional to the moment squared, as described in detail, for example, in Ref. 10. If the dipole moment is the leading term, as it generally is for a dipolar molecule, then stabilization is proportional to the dipole squared. If the dipole moment of the ground state is greater than that of the excited state under consideration there will be a blue shift—a shift to greater transition energy—as the ground state is more stabilized than is the excited state. The opposite is true if the excited state dipole is greater than that of the ground state. The larger the change in dipole moment, the larger the shift. If the dipole moment is zero, then the leading term would be the quadrupole squared, and a much smaller shift is expected.

Dispersion shifts are more subtle, and are due to the induced-dipole induced-dipole interactions of the chromophore with the solvent. Liptay has shown these are red shifts, as the excited state is always more ''dispersed'' than is the ground state.[41] One might expect that the higher the transition energy of the state, the greater the dispersion interaction and the greater the red shift. In general this is true,[36] and might be related to the greater polarizability of the higher excited states, as they generally have looser held electrons as seen from the perturbation theory expressions for the dispersion energy.[36,41] Other things that lead to greater dispersion are lower lying excited states in the solvent that can lead to a more polarizable solvent, and a smaller molecule, which allows for a closer contact between the solute and solvent molecule. These contacts enter into the $1/r^6$ dependence in the perturbation theory interaction for induced-dipole induced-dipole interactions. The playoff of these two effects, low lying available states, with benzene as

the solvent, and smaller size, with $H_2O$ as the solvent, explains why the dispersive red shift of benzene in these solvents, Table IV, only differ by a factor of 2.5 in spite of the fact that benzene is far more polarizable than water.

The present results should also help directly in calibrating such simpler models for dispersion as that described in Ref. 36, that rely on cavity radii. The present model does not have parametric dependence, but does require stochastic sampling, something accomplished indirectly using reaction field methods.

The present model has been used in a study of the polarizability of acetone in water,[42,43] a polar molecule in a polar solvent, with very good success. We are now examining this model for estimating the spectroscopic shifts of polar systems, such as acetone and pyridine, where much experimental information is also available. Since dipolar interactions decay as $1/r^3$ vs dispersion, $1/r^6$, we might expect a greater sensitivity to truncation of the cluster size for the polar systems. For these studies, the entire cluster will be surrounded by a reaction field. Dipole reaction field effects for an isolated molecule, such as pyridine in benzene, are proportional to $1/a_0^3$, where $a_0$ is the cavity radius. $a_0$ is approximately 3.2 Å for pyridine, and this cavity radius gives good results for pyridine in various solvents.[10] For a cluster of pyridine surrounded by 13 benzenes (the same as benzene surrounded by 13 benzenes in this study) these effects will be screened and reduced by a factor of at least 14 ($a_0^3$ is proportional to the volume and this is now 14 times larger). This factor should be an upper-bound in most cases as the first solvation shell will also polarize to reduce the net moments of the chromophore that are felt outside this shell.

Preliminary results on pyridine and acetone[42,43] and formaldehyde[44] are very good. The supermolecule approach described here cannot only account for the electrostatic terms and dispersion, but also for actual electron transfer between solvent and solute,[45] cases in which the solute and solvent are far from weakly interacting. The quantum mechanics required over the cluster is indeed considerably larger than that required over the chromophore alone, but it is certainly doable using semiempirical quantum mechanics such as that described here.

## ACKNOWLEDGMENTS

This work is supported in part by grants from CNPq and FAPESP and from the Office of Naval Research.

[1] C. Reichardt, *Solvent Effects in Organic Chemistry* (Verlag Chemie, Weinheim, New York, 1979).
[2] C. J. Cramer and D. G. Truhlar, in *Reviews in Computational Chemistry*, edited by D. B. Boyd and K. B. Lipkowitz (VCH, New York, 1995), Vol. 6.
[3] S. Basu, Adv. Quantum Chem. **1**, 145 (1964).
[4] A. T. Amos and B. L. Burrows, Adv. Quantum Chem. **7**, 289 (1973).
[5] M. Karelson and G. H. F. Diercksen, in *Problem Solving in Computational Molecular Science, Molecules in Different Environments*, NATO ASI Series, edited by S. Wilson and G. H. F. Diercksen (Kluwer, Dordrecht, 1997).
[6] N. S. Bayliss and E. G. McRae, J. Phys. Chem. **58**, 1002 (1954).
[7] E. G. McRae, J. Phys. Chem. **61**, 562 (1957).
[8] N. S. Bayliss, J. Chem. Phys. **18**, 292 (1950).
[9] O. Tapia and O. Goscinski, Mol. Phys. **29**, 1653 (1975).
[10] M. M. Karelson and M. C. Zerner, J. Phys. Chem. **96**, 6949 (1992), and references therein.
[11] A. Warshel and M. Levitt, J. Mol. Biol. **103**, 227 (1976).
[12] M. J. Field, P. A. Bash, and M. Karplus, J. Comput. Chem. **11**, 700 (1990).
[13] J. Gao and X. Xia, Science **258**, 631 (1992).
[14] D. Bakowies and W. Thiel, J. Phys. Chem. **100**, 10580 (1996).
[15] M. A. Thompson, J. Phys. Chem. **100**, 14494 (1996).
[16] (a) K. Coutinho, S. Canuto, and M. C. Zerner, Int. J. Quantum Chem. **65**, 885 (1997); (b) K. Coutinho and S. Canuto, Adv. Quantum Chem. **28**, 89 (1997).
[17] K. Coutinho, M. J. de Oliveira, and S. Canuto, Int. J. Quantum Chem. **66**, 249 (1998).
[18] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[19] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Clarendon, Oxford, 1987).
[20] K. Coutinho and S. Canuto, DICE: A Monte Carlo program for molecular liquid simulation, University of São Paulo, Brazil, 1997.
[21] W. L. Jorgensen, J. Chandrasekhar, and J. D. Madura, J. Am. Chem. Soc. **106**, 6638 (1984).
[22] I. R. McDonald, D. G. Bounds, and M. L. Klein, Mol. Phys. **45**, 521 (1982).
[23] W. L. Jorgensen, J. Chandrasekhar, and J. D. Madura, J. Chem. Phys. **79**, 926 (1983).
[24] G. M. Pearl, M. C. Zerner, A. Broo, and J. McKelvey, J. Comput. Chem. **19**, 781 (1998).
[25] M. Squilacote, R. S. Sheridan, O. L. Chapman, and F. A. L. Anet, J. Am. Chem. Soc. **97**, 3244 (1975).
[26] C. Chatfield, *The Analysis of Time Series: An Introduction*, 3rd ed. (Chapman and Hall, London, 1984).
[27] R. Krätschmer, K. Binder, and D. Stauffer, J. Stat. Phys. **15**, 267 (1976).
[28] R. Friedberg and J. E. Cameron, J. Chem. Phys. **52**, 6049 (1970).
[29] J. Ridley and M. C. Zerner, Theor. Chim. Acta **32**, 111 (1973).
[30] A. H. Narten, J. Chem. Phys. **48**, 1630 (1968).
[31] A. H. Narten, J. Chem. Phys. **67**, 2102 (1977).
[32] S. Tang and D. P. Landau, Phys. Rev. B **36**, 567 (1987).
[33] M. C. Zerner, ZINDO: A semiempirical program package, University of Florida, Gainesville, Florida 32611.
[34] M. C. Zerner, in *Reviews of Computational Chemistry*, edited by K. B. Lipkowitz and D. B. Boyd (VCH, New York, 1991), Vol. 2, pp. 313–366.
[35] N. S. Bayliss and L. Hulme, Aust. J. Chem. **6**, 257 (1953).
[36] N. Rösch and M. C. Zerner, J. Phys. Chem. **98**, 5817 (1994).
[37] F. M. Floris, J. Tomasi, and J. L. Pascual-Ahuir, J. Comput. Chem. **12**, 784 (1991).
[38] J. Tomasi and M. Persico, Chem. Rev. **94**, 2027 (1994).
[39] S. Canuto, K. Coutinho, and M. C. Zerner, J. Chem. Phys. **112**, 7293 (2000).
[40] J. Li, C. Cramer, and D. G. Truhlar, Int. J. Quantum Chem. **77**, 264 (2000).
[41] W. Liptay, Z. Naturforsch. **112**, 9405 (1965).
[42] K. Coutinho, N. Saavedra, and S. Canuto, J. Mol. Struct.: THEOCHEM **466**, 69 (1999).
[43] K. Coutinho, S. Canuto, and M. C. Zerner (work in progress).
[44] S. Canuto and K. Coutinho, Int. J. Quantum Chem. **77**, 192 (2000).
[45] G. Pearl and M. C. Zerner, J. Am. Chem. Soc. **121**, 399 (1999).